Patent Application
Attorney Docket No.: 14357RO

CONTROL PLANE FAILURE RECOVERY IN A NETWORK

CROSS-REFERENCE TO RELATED APPLICATIONS

A claim of priority is made to U.S. Provisional Patent Application No. 60/279927, entitled Distributed Recovery Method for Control Plane Failures in LDP, filed March 29, 2001, which is incorporated by reference.

FIELD OF THE INVENTION

The present invention relates generally to network communications, and more particularly to recovery from control plane failure in a communications network.

BACKGROUND OF THE INVENTION

Techniques for protecting and restoring services from network failures are well known. The techniques typically focus on failures in the data plane, although in some conventional networks the distinction makes little difference. SONET/SDH network failures, for example, necessarily imply simultaneous control plane

Express Mail Label No. EE430633152US, Date of Mailing 06/26/2001

and data plane failure because control messages and user information are transmitted together in frames. However, recently developed optical network architectures having separate control and data planes present difficulties. Known protection and restoration techniques may be employed with such optical networks to recover from faults in the data plane. However, these techniques appear to be less useful against control plane failure. It would therefore be desirable to have protection and restoration capability for the control plane.

SUMMARY OF THE INVENTION

In accordance with the present invention, selected control plane information is mirrored. The mirrored control plane information may be employed to facilitate restoration of the control plane in the event of failure. In one embodiment, connectivity resource information used by a first device is mirrored on a second device. For example, the information reflects

the available channels of the link between the first device and the second device. The devices may be Label Switched Routers ("LSRs"), and the label information may be label tables employed in Multi Protocol Label Switching ("MPLS"). In particular, a copy of the label information database in a downstream LSR is stored on an upstream LSR. Label information that is lost to the downstream LSR can then be recovered from the upstream LSR in the event of control plane failure.

The present invention is advantageously scalable. The label information in each mirror includes only information regarding a specific link. In particular, a mirror in an upstream LSR includes label information only from a downstream LSR label database, although the upstream LSR may maintain multiple mirrors in order to support multiple downstream LSRs. Redundant storage of label information for the network is therefore distributed, thus enabling the network to be more easily scaled than would be possible with a dedicated label information storage device.

The present invention will now be described in more detail with reference to exemplary embodiments thereof as shown in the appended drawings. While the present invention is described below with reference to preferred embodiments, it should be understood that the present invention is not limited thereto. Those of ordinary skill in the art having access to the teachings herein will recognize additional implementations, modifications, and embodiments, as well as other fields of use, which are within the scope of the present invention as disclosed and claimed herein, and with respect to which the present invention could be of significant utility.

BRIEF DESCRIPTION OF THE DRAWINGS

In order to facilitate a fuller understanding of the present invention, reference is now made to the appended drawings. These drawings should not be construed as limiting the present invention, but are intended to be exemplary only.

Figure 1 is a diagram of a portion of a network including a plurality of LSRs having Label Information Mirrors.

Figure 2 illustrates control plane link failure in the network of Figure 1.

Figure 3 illustrates control plane node failure in the network of Figure 1.

Figure 4 illustrates the LSRs of Figure 1 in greater detail.

DETAILED DESCRIPTION OF EXEMPLARY EMBODIMENTS

Figure 1 illustrates a portion of an optical network. The network includes nodes such as Label Switched Routers ("LSRs") 10, 12, 14, 16, and links 38, 40, 41, containing one or multiple channels. The LSRs employ Multi-Protocol Label Switching ("MPLS") to facilitate routing of traffic in Label Switched Paths ("LSPs"). Labels are distributed among the LSRs by employing a Label Distribution Protocol ("LDP"), such as constraint based routing LDP ("CR-LDP"). The LDP

runs in the control plane of the network, which is physically separate from the data plane. In the illustrated example, the control plane runs over an Ethernet network 18, while the data plane runs over a wavelength routed Dense Wavelength Division Multiplexing ("DWDM") network (not illustrated).

Logically adjacent LSRs, such as LSR 12 and LSR 14, communicate to establish common label information for transmitting traffic between them. Label information is exchanged between logically adjacent LSRs in LDP sessions 20, 22, 24. For example, label information is exchanged between LSR 12 and LSR 14 in LDP session 22. Each LDP session may include a plurality of LDP messages selected from four general types: Discovery messages, Advertisement messages, Session messages, and Notification messages. Discovery messages provide a mechanism whereby LSRs indicate their presence in a network, such as by periodically sending a "Hello" message. "Hello" messages are transmitted via UDP to the LDP port at the "all routers on this subnet" group multicast address. An LDP

initialization procedure is employed via TCP transport to establish an LDP session with an LSR discovered via a "Hello" type discovery message. Upon successful completion of the initialization procedure, the two LSRs become LDP peers, and may exchange advertisement messages. Advertisement messages, such as requesting a label or advertising a label mapping to a peer, are transmitted by an LSR based on local decisions made by that LSR. In general, an LSR requests a label mapping from a neighboring LSR when needed, and advertises a label mapping to a neighboring LSR when it is determined that use of the label by the neighbor may be advantageous. Session messages are employed for the actual exchange of label information. Notification messages are employed to notify network events. For example, they are to notify peers about newly available / unavailable channels in the data plane.

Each LSR maintains a label information database ("LID") containing the label information employed by that LSR. In the illustrated example, LSR 12 includes label information database 26, LSR 14 includes label

information database 28, and LSR 16 includes label information database 30. Each label information database includes mappings of labels associated with particular Label Switched Paths ("LSPs").

Label Information Mirrors ("LIMs") 32, 34, 36 are created in upstream LSRs to facilitate LDP recovery in the event of control plane failure. Each LIM is a copy of a label information database in a logically adjacent downstream LSR made via an LDP session. For example, LIM 32 in LSR 10 contains the label information from database 26, LIM 34 in LSR 12 contains the label information from database 28, and LIM 36 in LSR 14 contains the label information from database 30. The LIMs may be employed to restore label information that is lost or corrupted in the corresponding label information database.

The LIMs are initialized contemporaneous with LDP session initialization. Type-Length-Value objects ("TLVs") may be employed in the LDP Session Initialization message to facilitate initialization. In particular, an Advertisement of Mirror of Label

Information TLV, and an Advertisement of Label Information Database TLV may be employed. The Advertisement of Mirror of Label Information TLV is operative to notify the downstream LSR peer about the contents of the LIM stored in the upstream LSR. The Advertisement of Label Information database TLV is operative to notify the upstream LSR peer about the contents of the label information database stored in the downstream LSR.

Following initialization, each LIM is synchronized with the corresponding downstream label information database. When LSR 14 assigns a label, upstream LSR 12 updates its corresponding LIM 34 after receiving a LDP Label Mapping message from the downstream LSR peer 14. Consequently, both the LIM 34 and the label information database 28 are synchronized after the LSP setting up phase. In the LSP tearing down phase, the upstream LSR 12 updates its LIM 34 when it sends an LDP Label Release message to the downstream LSR peer 14. In this way, both the LIM 34 and the label information database 28 are synchronized after the LSP tearing down phase.

The upstream LSR 12 also maintains the LIM 34 when it receives an LDP Notification message from the downstream LSR peer 14. Consequently, in any stable state of the LDP operation, the LIM 34 and the label information database 28 are synchronized.

Figure 2 illustrates control plane failure and recovery from link failure. When the control channel failure occurs between LSR 12 and LSR 14, the LDP session 20 closes. Neither the LIM 34 (Fig. 1) residing in LSR 12 nor the label information database 28 (Fig. 1) residing in LSR 14 are emptied. After the failure of the control channel is repaired, LSR 12 re-initializes the LDP session with LSR 14. The contents of the LIM in LSR 12 are transmitted to LSR 14 in an LDP Session Initialization message. LSR 14 compares the contents of the LIM to its own label information database and calculates the intersection. LSR 14 then updates its label information database as the intersection calculated. LSR 14 then transmits the calculated intersection back to LSR 12 in an Advertisement of Label Information Database TLV. Upon

receiving the Advertisement of Label Information Database TLV in the LDP Session Initialization message from LSR 14, LSR 12 updates its LIM as the TLV indicates. Hence, the label information is recovered and the user data communication is not be interrupted.

Generally, the mirror in LSR 12 and the label information database in LSR 14 are synchronized in any stable state of LDP operation prior to control channel failure. When LSR 14 calculates the intersection of its own label information database and the advertised mirror of label information of LSR 12, the result is identical to either of them. However, when the control channel failure happens before the LDP operation reaches a stable state, their contents could be slightly different.

Referring to Figures 1 and 3, when a LSR is reset and re-initialized, the LIM and the label information database stored in it are set according to actual network configuration. After a single control node failure at LSR 12, the following recovery steps facilitate reset and re-initialized:

1. LSR 12 initializes both its LIM 34 and its label information database 26 according to actual network configuration.
2. LSR 12 sends an LDP Session Discovery message ("Hello" message) to LSR 10 through the link 38 between them.
3. LSR 10 advises LSR 12 about the contents of its mirror 32 in the LDP Session Initialization message.
4. LSR 12 compares the contents of the mirror 32 with its own label information database 26 and calculates the intersection.
5. LSR 12 updates its label information database 26 as the calculated intersection.
6. LSR 12 sends the intersection back to LSR 10 in an Advertisement of Label Information Database TLV.
7. Upon receiving the Advertisement of Label Information Database TLV in the LDP Session Initialization message from LSR 12, LSR 10 updates its LIM 32 as the TLV indicates, at which point the

label information database 26 in LSR 12 is recovered and synchronized with its mirror 32.

In parallel with steps 2-7, LSR 12 also recovers its LIM 34 corresponding to its outgoing links. In particular:

1. LSR 12 sends an LDP Session Discovery message ("Hello" message) to the LSR 14 through the link 40 between them.
2. LSR 12 advised LSR 14 about the contents of its LIM 34 in the LDP Session Initialization message. (The contents of its mirror of label information matches the actual link configuration between them.)
3. LSR 14 compares the contents of the LIM 34 with its own label information database 28 and calculates the intersection.
4. LSR 14 updates its label information database 28 as the calculated intersection. (This should not affect the contents of its label information database, because the advertised LIM 34 is a superset of its label information database.)

5. LSR 14 sends the intersection back to LSR 12 in an Advertisement of Label Information Database TLV.

6. Upon receiving the Advertisement of Label Information Database TLV in the LDP Session Initialization message from LSR 14, LSR 12 updates its mirror of label information as the TLV indicates. (Now the mirror of label information in LSR 12 is recovered and synchronized with its counterpart.)

Figure 4 illustrates LSR 12 (Fig. 1) in greater detail. LSR 12 has one incoming link 42 and one outgoing link 44. There is one downstream side LDP entity 46 corresponding to each incoming link. And each downstream side LDP entity 46 has a private label information database 48. Consequently, label information is stored on a per link basis. Similarly, there is one upstream side LDP entity 50 corresponding to each outgoing link 44. And each upstream side LDP entity 50 has it own mirror of label information (LIM) 52. Consequently, the label information database 48 has meaning with respect to only one LDP entity 46 and one link 42. Similarly, the mirror of label information 52

has meaning with respect to only one LDP entity 50 and one link 44. Different LDP entities in one LSR may employ different TCP/UDP port numbers to communicate with peers.

In an alternative embodiment, the present invention could be employed with a centralized recovery mechanism. The centralized recovery mechanism could be beneficial for inter-domain recovery. More particularly, if the control plane of a whole domain fails, and some kind of centralized control plane backup is provided for that domain, the control information can be recovered through a centralized method.

The present invention may also be employed to facilitate protection and restoration of control channels by using backup control channels. If recovery from failures in the control plane can be accomplished by using backup control channels, LDP sessions will not sense the failures and will not take any recovery action. Recovery by backup control channels may be preferable, in general. If the control plane cannot be

recovered within a time bound, the TCP and LDP layer would then take action.

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the present invention, in addition to those described herein, will be apparent to those of ordinary skill in the art from the foregoing description and accompanying drawings. Thus, such modifications are intended to fall within the scope of the following appended claims. Further, although the present invention has been described herein in the context of a particular implementation in a particular environment for a particular purpose, those of ordinary skill in the art will recognize that its usefulness is not limited thereto and that the present invention can be beneficially implemented in any number of environments for any number of purposes. Accordingly, the claims set forth below should be construed in view of the full breadth and spirit of the present invention as disclosed herein.